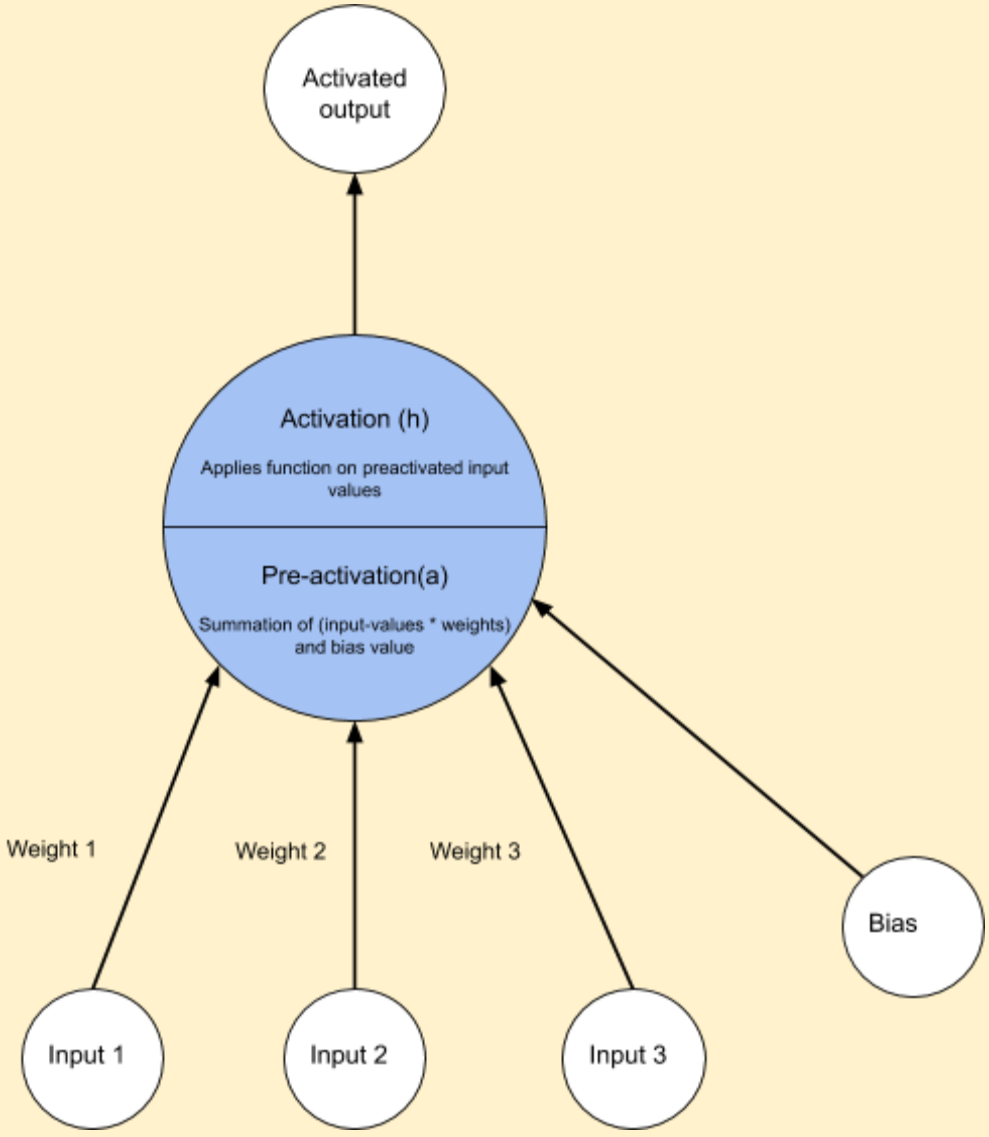
PadhAI: Deep Neural Networks

One Fourth Labs

A Generic Deep Neural Network

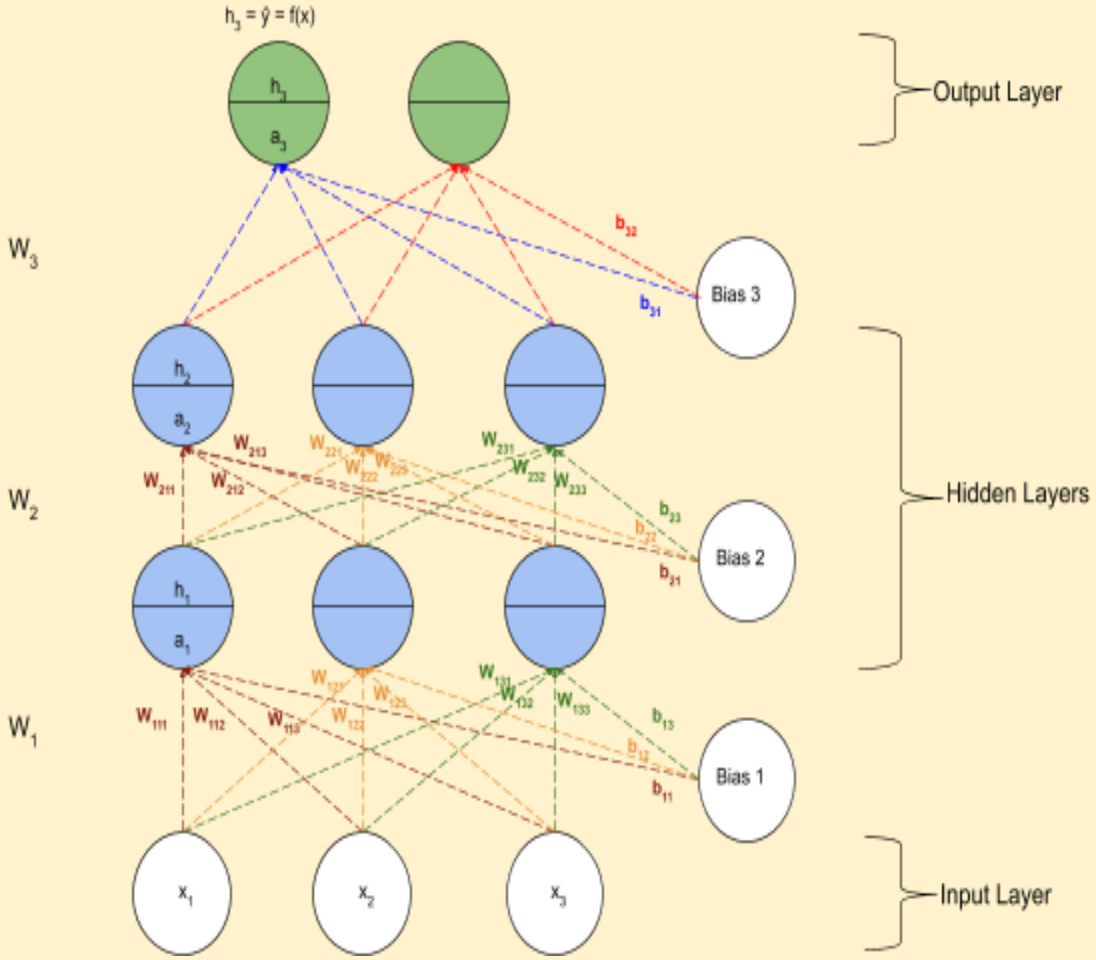
Can we clarify the terminology used?

1. Let us revisit the structure of a neuron



2. Let us break down the terms
   a. Let i refer to the layer being referenced
   b. **Pre-activation function** $a_i = \Sigma(input * weights) + bias$
   c. **Activation function** $h_i = \frac{1}{1+e^{-(a_i)}}$ a
   d. Here, the activation function is the sigmoid function.
   e. The construction of a Neural network is a simple stacking of these neurons in layers, one on top of the other
   f. The outputs of one layer of neurons become the inputs for the next layer.
   g. The cycle of pre-activation and activation repeats itself from the input layer, till we reach the output layer and obtain the desired function

3. Let us break down the structure of a Neural Network



4. Let's break down some of the terms used:
   a. The format of w is $W_{\text{(Layer number)(Next layer Neuron)(Current Layer Input/neuron)}}$
      i. So W_{213} refers to the weight corresponding to the 3rd input on 1st neuron of the 2nd hidden layer
   b. For each layer i where 0 <= i <= L
      i. Pre-activation $a_i(x) = W_i h_{i-1}(x) + b_i$
      ii. Activation $h_i(x) = g(a_i(x))$ where 'g' is called the activation function
      iii. Activation at output layer L is $f(x) = h_L = O(a_L)$ where 'O' is called the output activation function